Theory I: Database Foundations

07.2012

Languages

Paradigms

- Relational algebra
- Relational calculus
- SQL: not explicitly considered in this theory course!

Relational Algebra

Basic Operators

- delete attributes: Projection.
- select tuples: Selection.
- combine relations: Join.
- set operators: Union, Difference.

Projection

Student

MatrId	Name	Address	Semester
1223	Hans Eifrig	Seeweg 20	2
3434	Lisa Lustig	Bergstraße 11	4
1234	Maria Gut	Am Bächle 1	2

$\Downarrow$

Student'

MatrId	Name
1223	Hans Eifrig
3434	Lisa Lustig
1234	Maria Gut

Projection on tuples

- Let $R(X)$ be a schema, where $X = \{A_1, \ldots, A_k\}$.
- Let Y be a set of attributes, where $\emptyset \subset Y \subseteq X$.
- Let $\mu \in \mathrm{Tup}(X)$ be a tuple over X.
- $\mu[Y]$ is called projection of μ to Y:

$$\mu[Y] \in \mathrm{Tup}(Y),$$

$\mu[Y](A) = \mu(A), A \in Y$.

Projection on relations

- Let $r \subseteq \text{Tup}(X)$ a relation and $Y \subseteq X$.
- $\pi[Y]r$ is called projection of r to Y:

$$\pi[Y]r = \{\mu \in \text{Tup}(Y) \mid \exists \mu' \in r, \text{such that } \mu = \mu'[Y]\}.$$

Example

$r =$

A	B	C
a	b	c
a	a	c
c	b	d

$\pi[A, C](r) =$

Selection

Course

CourseId	Institute	Title	Description
K010	DBIS	Databases	Foundations of Databases
K011	DBIS	Information Systems	Foundations of Information Systems
K100	MST	Microsystems	Foundations of Microsystems

$\Downarrow$

Course'

CourseId	Institut	Title	Description
K100	MST	Microsystems	Foundations of Microsystems

Selection condition

- Let $A, B \in X$, $a \in dom(A)$, and $\theta \in \{=, \neq, \leq, <, \geq, >\}$ a comparison operator.
- An (atomic) selection condition α (on X) is of the form $A\,\theta\,B$, resp. $A\,\theta\,a$, resp. $a\,\theta\,A$.
- A tuple $\mu \in \text{Tup}(X)$ fulfills a selection condition α, if $\mu(A)\,\theta\,\mu(B)$, resp. $\mu(A)\,\theta\,a$, resp. $a\,\theta\,\mu(A)$ hold.
- Atomic selection conditions can be generalized to formulas using $\wedge$, $\vee$, $\neg$, and $(,)$.

Example

$$X = \{A, B, C\}.$$
$$\mu_1 = (A \rightarrow 2, B \rightarrow 2, C \rightarrow 1),\ \mu_2 = (A \rightarrow 2, B \rightarrow 3, C \rightarrow 2)$$
$$\alpha_1 = (A = B),\ \alpha_2 = ((B > 1) \wedge (C > 1))$$

Selection

- Let $r \subseteq \mathrm{Tup}(X)$ be a relation and α a selection condition over X.
- $\sigma[\alpha]r$ is called selection of relation r by α:

$$\sigma[\alpha]r = \{\mu \in \mathrm{Tup}(X) \mid \mu \in r \wedge \mu \text{ fulfills } \alpha\}.$$

Example

$r =$

A	B	C
a	b	c
d	a	f
c	b	d

$\sigma[B = b](r) =$

Union and difference

- Let X be a set of attributes and $r \subseteq \text{Tup}(X), s \subseteq \text{Tup}(X)$ two relations.
- $$r \cup s = \{\mu \in \text{Tup}(X) \mid \mu \in r \vee \mu \in s\}.$$ $$r - s = \{\mu \in \text{Tup}(X) \mid \mu \in r, \text{where } \mu \notin s\}.$$

Example

$r =$

A	B	C
a	b	c
d	a	f
c	b	d

$s =$

A	B	C
b	g	a
d	a	f

$r \cup s =$

$r =$

A	B	C
a	b	c
d	a	f
c	b	d

$s =$

A	B	C
b	g	a
d	a	f

$r - s =$

Join

Student

MatrId	Name
1223	Hans Eifrig
3434	Lisa Lustig
1234	Maria Gut

Registration

MatrId	CourseId
1223	K010
1234	K010

Course

CourseId	Title
K010	Databases
K011	Information System
K100	Microsystems

⇓

Student'

MatrId	Name	CourseId	Title
1223	Hans Eifrig	K010	Databases
1234	Maria Gut	K010	Databases

Join

- For sets of attributes X, Y, we may also write XY instead of $X \cup Y$.
- Let $r \subseteq \mathrm{Tup}(X), s \subseteq \mathrm{Tup}(Y)$.
- The (natural) join $\bowtie$ of r and s is defined:

$$r \bowtie s = \{\mu \in \mathrm{Tup}(XY) \mid \mu[X] \in r \wedge \mu[Y] \in s\}.$$

Example

$r =$

A	B	C
1	2	3
4	5	6
7	8	6

$s =$

C	D
3	1
6	2
4	5

$r \bowtie s =$

Observation about join

If $X_1 \cap X_2 = \emptyset$, then $r_1 \bowtie r_2 = r_1 \times r_2$.

Generalization of join

Let X_i, $1 \leq i \leq n$ be sets of attributes.

$$\bowtie_{i=1}^{n} r_i = \{\mu \in \mathrm{Tup}(\cup_{i=1}^{n} X_i) \mid \mu[X_i] \in r_i, 1 \leq i \leq n\}.$$

Renaming

- Let $X = \{A_1, \ldots, A_k\}$, $Y = \{B_1, \ldots, B_k\}$ be sets of attributes.
- Let δ be a bijection from X to Y, where $dom(A) = dom(\delta(A))$. If $\delta(A) = B$, we write $A \to B$.
- Consider relation $r \subseteq \text{Tup}(X)$.
- The renaming $\delta[X, Y]$ with respect to r is given as follows:

$$\delta[X, Y]r = \{\mu \in \text{Tup}(Y) \mid \begin{array}{l} \exists \mu' \in r, \text{such that} \\ \mu'(A_i) = \mu(\delta(A_i)), 1 \leq i \leq k\} \end{array}$$

Example

$X = \{A, B, C\}, Y = \{D, E, C\}$ und $\delta = \{A \to D, B \to E, C \to C\}$.

$r =$

A	B	C
a	b	c
d	a	f
c	b	d

$\delta[X, Y]r =$

Basic Operators

- Selection, projection, union, difference, and join are the basic operators of relational algebra.
- The valid expressions of the relational algebra can be defined inductively.
- We could define other useful operators.

further operators

Let X_i, $1 \leq i \leq n$, be formats and let $r_i \subseteq \text{Tup}(X_i)$, $1 \leq i \leq n$, be relations.

- *Intersection*. Sei $X_1 = X_2$.

$$r_1 \cap r_2 = r_1 - (r_1 - r_2).$$

- *θ-Join*. Let $X_1 \cap X_2 = \emptyset$ and let α b e an arbitrary selection condition over $X_1 \cup X_2$.

$$r \bowtie_\alpha s = \sigma[\alpha](r \times s).$$

If α uses only equality: *equi-join*.

Division

Let X_1, X_2 be formats, $X_2 \subset X_1$, $Z = X_1 - X_2$ and $r_2 \neq \emptyset$.

$$\begin{aligned} r_1 \div r_2 &= \{\mu \in \mathrm{Tup}(Z) \mid \{\mu\} \times r_2 \subseteq r_1\} \\ &= \pi[Z]r_1 - \pi[Z](((\pi[Z]r_1) \times r_2) - r_1). \end{aligned}$$

Example

$r_1 =$

A	B	C	D
a	b	c	d
a	b	e	f
b	c	e	f
e	d	c	d
e	d	e	f
a	b	d	d

$r_2 =$

C	D
c	d
e	f

$r_1 \div r_2 =$

Example

Course(<u>CourseId</u>, Institute, Name, Description)
Registration(<u>MatrId</u>, <u>CourseId</u>, Semester, Grade)

$\pi[\texttt{MatrId}](\texttt{Registration} \div \pi[\texttt{CourseId}]\texttt{Course})$

Algebra as a query language

- We cannot express all computable transformations over instances of database schemas.

  Example: transitive closure of binary relations.

Equivalence

Two algebra expressions Q, Q' are called equivalent, $Q \equiv Q'$, if for any instance $\mathcal{I}$ of a database:

$$\mathcal{I}(Q) = \mathcal{I}(Q').$$

Examples

Let $\text{attr}(\alpha)$ be the attributes in α and let $R, S, T \ldots$ be relation names whose formats are X, Y, Z.

- $Z \subseteq Y \subseteq X \implies \pi[Z](\pi[Y]R) \equiv \pi[Z]R$.
- $\text{attr}(\alpha) \subseteq Y \subseteq X \implies \pi[Y](\sigma[\alpha]R) \equiv \sigma[\alpha](\pi[Y]R)$.
- $R \bowtie R \equiv R$.
- $X = Y \implies R \cap S \equiv R \bowtie S$.
- $\text{attr}(\alpha) \subseteq X, \text{attr}(\alpha) \cap Y = \emptyset \implies \sigma[\alpha](R \bowtie S) \equiv (\sigma[\alpha]R) \bowtie S$.